Borradores de ECONOMÍA

How Productive is Rural Infrastructure? Evidence on Some Agricultural Crops in Colombia

Por: Ignacio Lozano-Espitia
Lina Ma. Ramírez-Villegas

Núm. 948
2016

BANCO DE LA REPÚBLICA · COLOMBIA
Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá

How Productive is Rural Infrastructure? Evidence on Some Agricultural Crops in Colombia

Ignacio Lozano-Espitia and Lina Ma. Ramírez-Villegas*

Abstract

This paper evaluates the role of rural infrastructure on the performance of some agricultural crops in Colombia. The study utilizes geo-referenced cross sectional data of four crops, coffee, rice, beans and plantains, collected for the majority of municipalities. Using genetic matching models, we find that both having access to irrigation and drainage systems and better infrastructure for marketing –rural roads and nearby retail and wholesale centers– significantly increase crop yield as well as planted and harvested areas. Results are robust to a suitable set of matching algorithms. The positive and significant impact on agricultural development provides support to reorient agricultural policy towards the supply of public goods that pushes up productivity.

Key Words: Public Goods; Agricultural Productivity; Irrigation System; Road Maintenance; Treatment Effect Models

JEL Classification: H41, Q12, Q15, R42, C21

* Senior researcher at *Banco de la República* (Central Bank of Colombia) ilozanes@banrep.gov.co and research intern (student of economics at *Pontificia Universidad Javeriana*) lramirezv@javeriana.edu.co, respectively. The opinions expressed here are those of the authors and do not necessarily represent either those of the Banco de la República or of its Board of Directors. As usual, all errors and omissions in this work are our responsibility

1. Introduction

Agriculture has been historically a priority in developing economies. Although its GDP contribution tends to diminish as countries reach higher levels of income, progress in agriculture will remain essential for economic growth, employment opportunities, poverty alleviation and, especially, food security. For Colombia, the latest figures show that agriculture contributes nearly 5% to the national GDP and 9.6% to its exports. Twenty five years ago, these figures reached up to 16% and 33%, respectively. The current public investment comprises 11% of the production value, even though its participation in the national budget is below 5%. In the labor field, however, agriculture is still absorbing nearly a fifth of the workforce (17.5%).

Progress in agriculture has been closely associated to rural infrastructure. Other factors that appear to affect agricultural development are related to land availability, market efficiency, quality of institutions, the use of improved seeds and fertilizers, and access to technology and credit. Nonetheless, deficiencies in physical infrastructure could not only reduce factor productivity and crop yields but also weaken market competitiveness and limit their spatial and temporal integration (Fan, Zhang and Rao, 2004; Pinstrup-Andersen and Shimokawa, 2006).

The infrastructure services for agriculture are diverse in nature. They comprehend but are not limited to land improvement programs (especially irrigation and drainage systems); road and railway networks; transportation systems; market centers and wholesale terminals with easy access to both transport services and financial intermediaries; electrical energy; and communication networks involving telephones, radios, and information-disclosure (Timmer, P, 2002). Because the majority of them may be classified as public goods (of collective consumption), in the sense of non-rivalry and non-exclusion, their provision must be supported or regulated by government. The technical assistance, research and development, and preferential access to credit and insurances also require government actions in order to remove market failures and create the conditions required by modern agriculture.

In this paper we pursue a twofold objective, to estimate the coverage of irrigation systems and infrastructure for marketing and also to assess their role on the performance of the main crops in Colombia. In this regard, this paper explores more deeply the issues addressed by Lozano and Restrepo (2016) which is the first attempt to provide empirical evidence on these subjects. We select irrigation and drainage systems because it is the leading land-improvement program and it is expected that water facilities have positive impacts on crop yield. Similarly, we choose the rural road network and wholesale center access facilities as the major marketing infrastructure, because their larger coverage could generate incentives to expand the agricultural frontier (assigning more land to planting and harvesting).

Given the endogeneity problem that may arise and the sample characteristics, we use *Genetic Matching* (*GM*) models as the empirical assessment technique. Because the provision of agricultural infrastructure is associated with the size and management of public budget, from the results we recommend to re-orientate agricultural policy towards the supply of public goods that push up productivity. This suggestion is crucial especially because of the recent debate on public resources distribution within the agricultural sector which was settled in favor of direct subsidies to producers and against investment in rural infrastructure.

Following this introduction, the paper is organized as follows: in Section 2 we present some trends on public budget distribution in the agricultural programs in Colombia. In Section 3 the coverage of infrastructure programs under study are estimated and analyzed. In Section 4 we estimate the *GM* models which allow to assess the impact of major infrastructure goods on crop yields and on the planted and harvested areas. Results are presented and discussed in section 5. The paper concludes with section 6.

2. Overview on Public Financing of Agriculture

The government expenditure to provide public goods for agriculture is annually assigned through the national investment budget. Historical data indicates that the overall budget of the Colombian government has fluctuated between 25% and 28% of the GDP since 2000 and more than half (52.5%, on average) has been employed for state functioning, 30% for debt service and less than one-fifth (18%) for investment. If the latter is broken down by sectors, we can see that agriculture has had a really marginal share into the investment budget (Table 1, panel 1). It represented 4.4%, on average, between 2003 and 2014. The sectors which have absorbed the majority of investment resources are social security and protection (29.7%), infrastructure (24.6%: mines, energy, transport and communications), human capital (12.3%) and defense, security and justice (8.4%).

Data on investment in agriculture looks slightly better when compared to its GDP (Table 1, panel 3). For 2014, it has recovered the level reached twenty years ago, when the sector achieved significant government compensation because of the trade liberalization policies (reached up to 13.3% in 1996). The lowest investment levels were recorded in the crisis of the early century (in 2000), falling to 2.3% of the agricultural GDP.

Table 1. Public Investment in Agriculture in Colombia
(Budget distribution for some selected years, percentages)

1. Relative to Other Sectors		**2002**	**2005**	**2008**	**2011**	**2014**	**Average**
Agriculture		4.2	3.2	6.2	5.0	5.0	4.4
Defense, Security and Justice		10.0	6.3	18.8	4.8	5.3	8.4
Human Capital		11.8	13.3	11.7	10.9	13.4	12.3
Infrastructure		14.9	19.4	21.1	27.0	26.9	24.6
State Operations		25.1	24.8	13.6	21.0	9.8	16.7
Social Security & Protection		31.4	30.7	24.8	25.6	35.0	29.7
Other		2.6	2.3	3.7	5.7	4.7	4.0
2. Distribution by Program Within Agriculture		**2000**	**2005**	**2008**	**2011**	**2014**	**Average**
Land Allocation		5.9	0.8	0.8	6.2	10.2	3.1
Competitiveness, Entrepreneurship & Others		4.2	3.8	3.0	3.7	3.4	4.1
Rural Housing		13.9	5.9	4.7	14.3	14.5	8.6
Direct Aids		4.1	5.0	53.0	36.8	16.6	25.1
Irrigation Systems		7.8	9.3	15.9	12.5	3.8	12.1
Credit Subsidies		36.1	11.0	6.8	14.9	11.6	14.4
Technical Assistance & Research		28.1	14.8	8.3	8.1	14.3	14.9
Price Stabilization Funds		0.0	49.4	7.6	3.4	25.7	17.6
3. Other Indicators	**1990**	**1996**	**2000**	**2005**	**2010**	**2014**	**Average**
Public Investment to Agriculture GDP Ratio	8.6	13.3	2.3	2.7	6.0	11.3	6.3
Agricultural to Total GDP Ratio	9.5	8.3	8.5	7.4	2.7	2.7	4.6

Source: Calculations by the authors with data from DNP –National Department of Planning–, and DANE –National Administrative Statistic Department–

To identify the major agriculture-related programs, we grouped the investment budget data executed by the Agriculture Ministry in eight broad categories (Table 1, panel 2). Clearly, the priority of agricultural programs has changed through time and so have the instruments to promote rural development. In the first half of the 2000s decade, the policy focused mainly on price support schemes (Price Stabilization Funds for rice, cotton, sugar, etc.), but the relevance of these programs disappeared during subsequent years. Credit subsidies and technical assistance were also dominant in this first period but not so in the following. Two thirds of the sector's budget was spent, on average, in these three programs.

The agricultural policy emphasis changed markedly since 2006. Direct aids to producers through the program *Agro Ingreso Seguro*, *AIS*, became the main support mechanism to farmers. To this direct aid program more than half of the public resources were assigned (53% in 2008). In theory, the AIS program sought to protect local producers against the competition that would create the free trade agreement (FTA) with the United States. In the last years, an important share of the public budget has been used to acquire and assign land to vulnerable population, to subsidize rural housing and to support technical assistance. However, the majority of resources were assigned to provide direct subsidies to coffee producers (through the PIC program in 2013-14), a pressure group that has historically shown high ability to lobby.

The direct subsidies to farmers have been amply criticized because they go against the principle of efficiency and equity that should guide the allocation of public funds (Steiner, et. al., 2015; Echavarría, 2014). In addition, with direct subsidies to a particular sector, the government sends wrong signals to the rest of farmers and, furthermore, reduces the available funds required to strengthen rural infrastructure. In fact, the leading program of land improvement (irrigation and drainage systems) has had a relatively marginal budget share along these years (12%, on average).

3. Rural Infrastructure

3.1 Systems of Irrigation and Drainage

Irrigation systems are the ideal mechanism of water regulation which, potentially, could incentive the land use for agriculture, improve their yield and facilitate the implementation of new technologies (Schoengold and Zilberman, 2007). From a labor market perspective, once the irrigation systems are implemented new land is incorporated into production, therefore engaging more workers in agribusiness. Likewise, food security has usually benefited from water irrigation programs both in volumes and in product quality (FAO, 1996).

International evidence on some of these attributes is conclusive: 40% of the food produced worldwide comes from land with irrigation systems and such production uses only 17% of the arable land. This implies that the yield per hectare of food benefited from irrigation far exceeds that of those that do not have such schemes, which is reflected finally in large differences in both value of production and total factor productivity (Dregne and Chou, 1992; Evenson, Pray and Rosegrant, 1998; Fan, Zhang and Rao, 2004; Magno, Cardoso and Salvato, 2008; Shenggen and Zhang, 2004).

In Colombia, there exist 512 irrigation districts of diverse scale (489 small, 9 medium and 14 large), even though one third of the smaller were not operating up to 2007. The scale is associated to the surface benefited (irrigated) of each one. The first irrigation systems were built in the nineteen thirties and thereafter, but especially in the seventies (Incoder, 2012). Currently, there are around 346 municipalities (from 1,121) accessing irrigation systems; even so, the benefited areas are relatively small.

To estimate the irrigation system coverage, we use data on irrigated area and data on potentially irrigable area with moderate biophysical constraints coming from the Agriculture Ministry (Unit of Rural Agricultural Planning, UPRA) and from SIGOT (Geographic Information System for Planning and Territorial Ordering), respectively. More precisely, coverage is defined as the area (in kms^2) effectively irrigated by each 1000 kms^2 of potentially irrigable area. At national level, Colombia reached a coverage ratio of 7,6% which is very low, especially when compared regionally (México 66,1% in 2009; Chile, 44,3% in 2003; Peru 40% in 2012; Brazil 18,4% in 2010 and Argentina 14,7% in 2011).

When examining irrigation coverage across municipalities, we found great discrepancies within regions. To illustrate, we group the coverage in three categories (Figure 1, panel A): relatively high coverage (with above 9 kms^2 irrigated by each 1000 kms^2, in red); medium coverage (between 2.5 and 9 kms^2 in orange) and low coverage (lower than 2.5 kms^2, in yellow). As it can be seen, the majority of municipalities do not have access to irrigation districts, simply because there are none available (no IS, in white) and only a very small number (178) has access to this type of infrastructure in restricted circumstances. The districts are concentrated in the center-east and center-west of the country and they benefit mainly 89 crops.

3.2. Marketing Infrastructure

To raise the productivity of rural activities an efficient marketing system is required. For such purpose, there must be a set of components interlinked into rural infrastructure, which includes roads to connect farms to markets, market centers and wholesale terminals with easy access to transport facilities, reliable supplies of electricity, communication networks, etc. We prioritize two components within marketing infrastructure which could have an important impact in the Colombian case: the rural road network and the retail and wholesale center access facilities which are used by farmers to commercialize.

a. The Rural Road Network

A wide network of good quality rural roads is crucial to expand the agricultural frontier and to accelerate rural development. In the rural environment, good quality roads imply, at least, that they are affirmed (if unpaved) and that they remain in good condition both in dry and rainy seasons. Because of positive effects expected on transaction costs, the access to a good road system could encourage expanding crop production since it facilitates the entry of inputs (including technical assistance) and allows the crop transportation to the marketing centers.

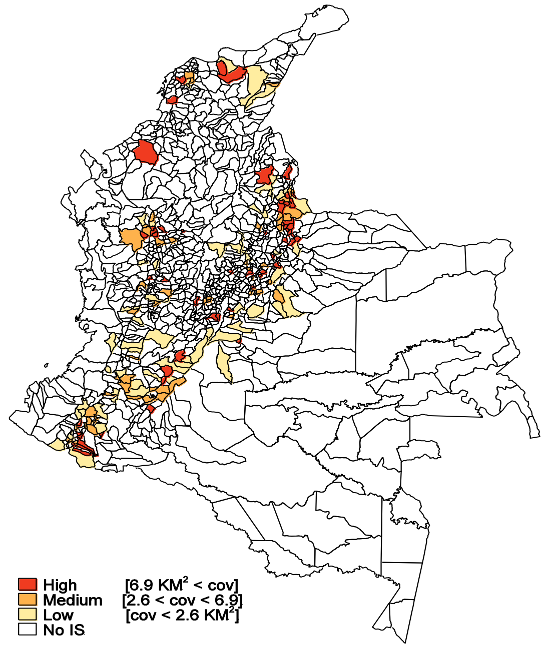
Colombia has large gaps in road infrastructure, especially when compared against emerging and developed countries. The delay in both coverage and quality of roads places the country at the highest cost per-tons transported by kilometer (OECD, 2015). Other transport options such as freight railroad and river systems are really marginal within the total tonnage that moves the country (around 15% and 4% respectively, according to Calderon and Servén, 2010). Heavy restrictions faced by farmers to transport their inputs and harvests are both by those roads that connect rural areas to municipalities (tertiary road network), as well as those leading to marketing centers and to ports (secondary and primary road networks).

Following standard definitions, we estimate the road coverage as its length per each 1000 km^2. The road length information comes from the Geographical Institute Agustin Codazzi, IGAC, which thereafter is categorized by the UPRA in six types. We call tertiary network the road types 4, 5 and 6 because they are the most closely associated with what we understand as rural infrastructure. The national network has 294,000 kms of roads, which yields a coverage of 41.2 kms. However, to make figures comparable with those released by the Ministry of Transport and international standards, it is necessary to subtract the roads of category 6 because they are narrow roads without affirming (trails) and only passable in dry weather. Thereby, the road network is reduced to 147,000 kms which implies a coverage rate of 22.9 kms. Within the total road network, only 12.7% are paved, which is similar to Peru (13%) and Brazil (13%), but really below Argentina (23%), Chile (23%) and Mexico (36%) (World Fact Book, 2015).

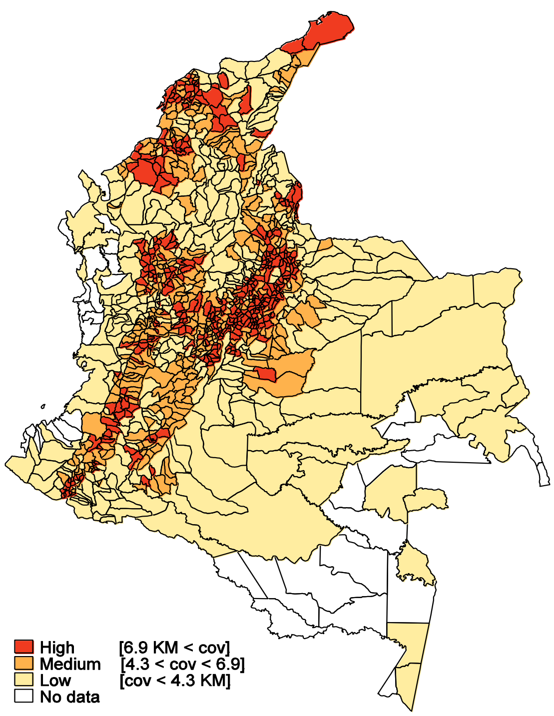
Figure 1 (panel B) provides evidence on the municipal differences in tertiary road network coverage. As in the previous case, we classify the coverage in three groups: municipalities with relatively high coverage (with above 7 kms of roads by each 1000 kms^2, in red); medium (between 4.3 and 7 kms, in orange) and low (lower than 4.3 kms, in yellow). Once again, the majority of high coverage municipalities are concentrated in the center-east and center-west of the country, even though municipalities near the Caribbean coast (to the north) have, also, relatively high coverage. Note that a large part of the country (75%) has poor coverage in rural roads (yellow area in the map) possibly affecting rural economies, whose income typically comes from crops marketed in urban areas. The low coverage of the road network could be associated to the low land use for agriculture, as previously mentioned.

Figure 1. Coverage of Rural Infrastructure in Colombia

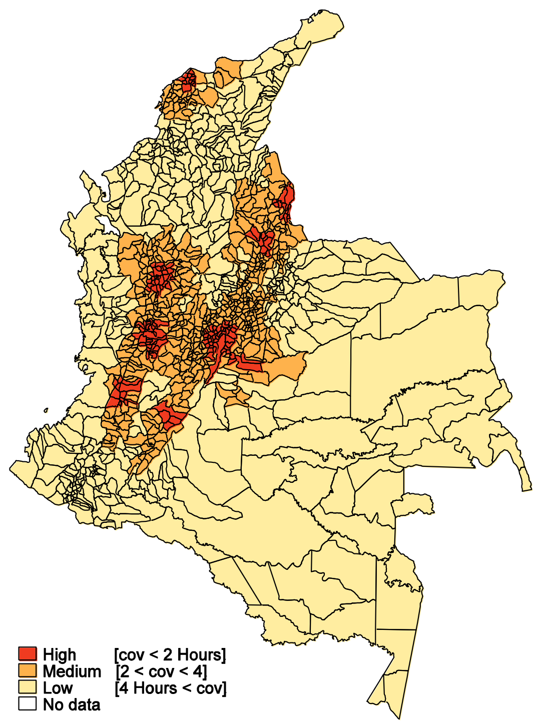
A. Irrigation and Drainage Systems[1]



B. Tertiary Road Network[2]



C. Isochrones to Marketing Centers [3]



[1] coverage (cov) = km² irrigated by each 1000km²of area potentially apt be irrigated

No IS: No irrigation systems available

[2] coverage (cov) = km of road constructed by each 100km²of rural area

[3] coverage (cov) = time spent by farmers to the nearest market (in hours)

Source: Calculations by the authors with data from UPRA

Indeed, the correlation between the areas under cultivation (as percentage of total area, in hectares) in 1081 municipalities, with respect to the total coverage of the road network (measured in kilometers of road per 1,000 hectares) for 2012 is positive (0.11) and statistically significant.

b. Access Facilities to Marketing Centers

The ease of access to retail and wholesale centers depends mainly of an adequate road network as well as the availability of nearby marketing centers. These two factors along with the average speed associated to road quality, are used to estimate the average time spent by farmers to get their crops to the nearest market. The resulting indicator, commonly known as *isochrones*, allows us to know how difficult (costly) is it for farmers to sell (or buy) their products (inputs). Because farmers of a big rural-municipality could be located at different distance (and travel time) from the nearest market, we use the dominant isochrones (i.e., travel time range which concentrates the majority of municipality area). The UPRA uses data from the Transport and Agriculture Ministries to construct the dominant isochrones to retail and wholesale centers, which we employ as an important proxy of market infrastructure.

Figure 2, panel C, shows the results. At national level, 95.7 million hectares (i.e., 84% of the total land area) are further away than a four-hour trip to the nearest marketing center and only 2.9% of rural areas are at less than a one-hour trip. The differences between regions are evident. We can see that municipalities with better marketing infrastructure are located at the center-east and center-west (coffee region) of the country with a time-trip of less than two hours (in red). Then there are those municipalities with travel times between two and four hours (in orange) and later the municipalities with greater than four hours' time (in yellow).

4. Methodological Framework

We employ the Genetic Matching (*GM*) technique, a generalized version of propensity score matching (*PSM*), to assess the role of rural infrastructure on the performance of five of the main crops of the Colombian agriculture. In particular, we are interested in assessing the effect of irrigation and drainage systems on crop yields and the effect of marketing infrastructure (rural roads and isochrones to retail and wholesale markets) upon planted and harvested areas. The former is tested because it is expected that permanent water access increases production per hectare, and the latter because better marketing infrastructure could translate into lower transaction costs, which, in turn, could encourage farmers to intensify their production processes via expanding the agricultural frontier.

The PSM model has become a popular method to assess the impact of public programs. Though widely used to evaluate labor market policies (LaLonde, 1986; Dehejia and Wahba, 1999; Heckman, Ichimura, and Todd, 1998; Peikes, Moreno and Orzol, 2008), it has been employed in other fields of study. The basic idea behind this technique is to compute the effect of a treatment variable (the program) on a particular outcome variable inherent in a group of individuals, conjecturing about how these individuals would have performed if they had not received the treatment (the unobservable or counterfactual outcome).

4.1 The Model

The model involves three pillars: individuals, treatment and potential outcomes. Following the standard notation for the case of a binary treatment, the treatment indicator D_i equals one if individual *i* receives treatment and zero otherwise. The potential outcomes are denoted as $Y_i(D_i)$ for each individual *i*, where *i = 1,..,N* and N denotes the total population (Caliendo and Kopeinig, 2005). The treatment effect for an individual *i* can be written as:

$$\tau_i = Y_i(1) - Y_i(0) \quad [1]$$

Note that if the person *i* was treated, it is not possible to estimate the individual treatment effect τi because the second term of the right-side of [1] is unobservable. Therefore, the counterfactual for each person must be derived somehow. Later, it will be shown how the counterfactual is found using the PSM technique. For this study we will focus on the average treatment effect on treated population (ATT). So, equation [1] can be rewritten as

$$\tau_{ATT} = E[\tau_i \,|D_i = 1] = E[Y_i(1)\,|D_i = 1] - E[Y_i(0)|D_i = 1] \quad [2]$$

where E[.|D] denotes the conditional expectations operator. We note that the simple OLS model representation of this effect would be given by $Y_i = \beta_0 + \tau_i D_i + \mu_i$, where $\tau_i = Y_i(1) - Y_i(0)$, and $Y_i(0) = E[Y_i(0)] + \mu_i = \beta_0 + \mu_i$. Revisiting the subject of the counterfactual term $[Y_i(0)|D_i = 1]$, it will be taken from individuals that did not receive the treatment, $E[Y_i(0)|D_i = 0]$, but are similar in characteristics to those individuals that did receive the treatment (Bernal and Peña, 2014).

Clearly, the substitute (counterfactual) term is appropriate if and only if

$$E[Y_i(0)\,|D_i = 1] - E[Y_i(0)|D_i = 0] = 0 \quad [3]$$

which implies that $E[\mu_i|D_i] = 0$; i.e., all individuals (treated and non-treated) should be identical in the set of observable characteristics (covariates), X, which are contained into μ_i. This is also known as the *conditional independence assumption*, CIA, meaning that given a set of observable covariates X which are not affected by treatment, potential outcomes are independent of treatment assignment.

In practice the observable characteristics (covariates), X, are used to match, which may carry computational problems in case of a high dimensional vector X. To deal with this dimensionality problem, Rosenbaum and Rubin (1983) suggest to use the so-called balancing scores. The propensity score, PS, i.e. the probability for an individual to participate in a treatment given his observed covariates X, $P[D = 1|X] = P[X]$, is one possible balancing score. Hence, the propensity score matching ATT estimator can be written as:

$$\tau_{ATT}^{PSM} = E_{P(X)|D=1}\{E[Y(1)|D = 1, P(X)] - E[Y(0)|D = 0, P(X)] \qquad [4]$$

Simply put, the PSM estimator is the mean difference in outcomes over the common support, appropriately weighted by the propensity score distribution of participants (Caliendo and Kopeinig, op., cit., 2005).

The last requirement to estimate equation [4] involves choosing a suitable matching algorithm to select the appropriate neighbors for each treated individual as well as weights assigned to these neighbors. Following literature suggestions, we will use the nearest neighbor (NN) matching, both with replacement (NNWR) and without replacement (NNNR) (Smith and Todd, 2005). Under this approach the individual from the comparison group is chosen as a matching partner for a treated individual that is closest in terms of propensity score. In the replacement case, an untreated individual can be used more than once as a match, whereas in the other case it is considered only once. Because the NN matching faces the risk of finding bad matches, we also use the caliper algorithm which imposes a tolerance level on the maximum propensity score distance. This way bad matches are avoided and hence the matching quality may rise.

It should be noted that GM is a matching algorithm that improves covariate balance, as opposed to other matching algorithms (Diamond and Sekhon, 2013). To do so, it constructs a modified Mahalanobis distance that accounts for different weights for each variable, and calculates the respective measure for each pair of observations (Diamond and Sekhon, 2013). Later it matches the observations using the desired method (NN, caliper) and the computed distance. This process is carried out with several weights at the same time, and over again until a convergence condition related to a loss function associated to each matched data set is met.

4.2 Data

Table 2 describes the variables employed for the estimates and the criteria used to turn the treatment binary. The treatment variables refer to the three main public infrastructure programs (irrigation and drainage systems, tertiary road network and retail and wholesale markets access facilities) while the outcome variables comprehend the yield and the planted and harvested areas of four crops: Coffee, Rice, Bean and Plantain. Planted and harvested areas are measured as the proportion of the total municipality area. These crops are important within the agricultural production of the country and their yield is measured as physical production (in tons) by hectare. Colombian municipalities play the role of individuals within the model. The geo-referenced data comes from UPRA for the years 2008 and 2013.

Table 2. Variables Employed in GM Estimates

Outcome Variables: Yield and Planted/Harvested Areas	Treatment Variables	Characteristics of Municipalities
Coffee Rice Bean Plantain	Irrigation and drainage systems = 1 if there exists at least one = 0 Otherwise Tertiary roads network = 1 if coverage is above the median = 0 Otherwise Isochrones to wholesale markets = 1 if travel time is above of the median = 0 Otherwise	Degree of agricultural potentiality; Land with potential for irrigation systems; Population (total and rural); Credit granted by FINAGRO

We use a reasonable set of characteristics for the municipalities which is restricted by data availability. The physical, demographic and financial characteristics, X, include the degree of agricultural potential (feature that captures the effect of biophysical characteristics), the land with potential for irrigation systems, total and rural population, and subsidized credit granted through Finagro (public agency specialized to support this sector). These characteristics are used for matching in four ways: firstly, using all variables jointly; secondly, adding the previous characteristics the square of each one; thirdly, adding to the first option all possible pairwise interactions between characteristics; and, finally, combining the three previous alternatives. Even though we estimate models with all matching alternatives described, we show in the results the models that achieve higher balance. Regarding the treatment variables, in Section 3 we provided details on the procedures followed to estimate the coverage of the rural infrastructure services and the sources of the data.

5. Results

Results about the incidence of irrigation systems on the yield of agricultural crops are show in Table 3. We select three of them only (rice, coffee and bean); nonetheless, results for the remaining models are available upon request. A couple of matching procedures are chosen for each crop, including nearest neighbor with replacement (NNWR) and Caliper at different standard deviation (SD) values (0.05; 0.1; 0.2; 0.5; 1).

Table 3. Impacts of Irrigation and Drainage Systems on Yield for Some Selected Crops Estimate by GM

YIELD	RICE				COFFEE				BEANS			
Matching Procedure	NNWR	CALIPER			NNWR	CALIPER			NNWR	CALIPER		
		0.2 SD	0.5 SD	1 SD		0.05 SD	0.1 SD	0.2 SD		0.05 SD	0.1 SD	0.2 SD
Total Average Effect	3.545	1.715	2.050	2.379	-0.110	0.029	0.045	0.130	0.035	0.106	0.104	0.257
P-Value	0.000	0.000	0.000	0.024	0.319	0.000	0.000	0.001	0.755	0.000	0.031	0.000
Sample	227	21	98	194	487	13	67	182	505	19	103	188
Treated	39	9	28	35	115	4	22	65	127	6	34	65
Non treated	188	12	70	159	372	9	45	117	378	13	69	123
Balance Test P-Val												
Agric. Potential.	0.259	0.176	0.001	0.038	0.275	0.150	0.325	0.363	0.235	1.000	0.583	0.024
I&D Potential	0.108	0.018	0.324	0.896	0.006	0.150	0.479	0.651	0.111	1.000	0.715	0.184
Rural Population	0.029	0.661	0.724	0.443	0.353	0.541	0.273	0.534	0.407	0.470	0.635	0.528
Total Population	0.022	0.270	0.414	0.120	0.156	0.965	0.114	0.779	0.227	0.291	0.550	0.829
Subsidized Credit	0.240	0.014	0.004	0.024	0.194	0.488	0.334	0.546	0.175	0.947	0.541	0.491
Controls												
Interactions	No	Yes	No	Yes	No	No	Yes	Yes	No	No	No	No
Squares	Yes	No	Yes	No	No	No	No	No	No	Yes	No	Yes

Note: SD stands for standard deviation. NNWR stands for nearest neighbor with replacement.

Four models are shown for each crop in order to prove the robustness of the results. These were chosen by the degree of balance they achieve based on the mean difference t-test between treated and non-treated. The average treat on the treated effect parameter (ATT), presented in the top line, captures the impact. Clearly, the effect is positive and statistically significant (below 0.05 levels) across all regressions, for the three crops presented.

In the case of rice, for instance, differences in the average yield (tons per hectare) between the produced in municipalities which benefit from irrigation systems and those that have no access, oscillates between 1.7 (Caliper 0.2 SD) and 3.5 (NNWR) tons per year. The sample covers 227 municipalities (39 treated and 188, non-treated, using

NNWR), and cross sectional estimates were made for 2008. In the middle section of Table 3, we also include the balance score tests for each of the variables used to match. These tests provide support for the required balance after the matching process. Finally, the way in which control variables are introduced in the matching algorithm (just the variables, their interactions, their squares or both) is reported at the bottom of the Table.

Results also confirm the positive and statistically significant effects of irrigation systems on the coffee and bean yield across all models. Nonetheless the size of the parameters is smaller, slightly below 0.13 and 0.26 (Caliper 0.2 SD), respectively. For these two models, note that the sample is smaller (especially for non-treated municipalities) and performance of the balance score tests is slightly better. The tradeoff between the sample size and results for balance scores is common when applying the GM technique.

Comparing results on these three crops, we highlight that the larger impact of irrigation systems on the rice yield could be associated with the higher water requirement for its growth cycle. In fact, land that is used for rice production needs high humidity levels permanently, which are ideally provided by irrigation systems. So, farmers who have access to such systems can plant two crops annually while those who do not have access are restricted only to one, due to the unimodal raining regime. The humidity requirement is relatively lower for coffee and bean; and hence, the smaller size of the impact.

Table 4 shows estimates on the incidence of one of the most relevant infrastructure tools for marketing (the rural road network) upon planted and harvested areas. For a given crop in one municipality, the planted and harvested areas may differ because diverse circumstances that arise along the production cycle; a negative weather shock, for instance, may reduce the harvested but not the planted area. Therefore it is advisable to make independent assessments. Once again we select areas of three products (rice, coffee and plantain) and the rest of results is available upon request. Likewise, we choose four models for each crop (NNWR and Caliper 0.1; 0.2; 0.5 and 1, SD) based on the balance achieved by each model.

We confirm anew the positive and statistically significant impact of better rural roads coverage on both planted and harvested areas across all regressions for the three crops considered. For the case of coffee, for example, the parameter associated to the total average effect indicates that the planted area in those municipalities with above the median tertiary road coverage, is greater in 2.9 percentage points compared to planted areas of municipalities that are below of the median (NNWR). The sample for these models is 486 municipalities (241 treated and 245 non-treated) and estimates were made with more recent data, from 2013.

Table 4. Impacts of Rural Roads on Planted and Harvested Areas for Some Selected Crops. Estimate by GM

PLANTED AREAS	RICE				COFFEE				PLANTAIN			
Matching Procedure	NNWR	CALIPER			NNWR	CALIPER			NNWR	CALIPER		
		0.1 SD	0.2 SD	0.5 SD		0.1 SD	0.5 SD	1 SD		0.2 SD	0.5 SD	1 SD
Total Average Effect	0.025	0.016	0.021	0.026	0.029	0.023	0.034	0.030	0.014	0.011	0.010	0.014
P-value	0.049	0.000	0.000	0.012	0.000	0.000	0.000	0.000	0.001	0.000	0.010	0.001
Sample	227	14	69	169	486	46	388	472	574	283	492	551
Treated	113	8	38	91	241	20	174	232	286	138	246	275
Non treated	114	6	31	78	245	26	214	240	288	145	246	276
Balance Test P-Values												
Agric. Potentiality	0.022	0.379	0.671	0.486	0.216	0.256	0.757	0.209	0.120	0.512	0.165	0.862
I&D Potential	0.103	0.277	0.478	0.178	0.884	0.551	0.188	0.209	0.939	0.329	0.222	0.236
Rural Population	0.522	0.124	0.932	0.133	0.967	0.280	0.203	0.497	0.985	0.342	0.477	0.235
Total Population	0.480	0.220	0.334	0.122	0.398	0.366	0.579	0.781	0.661	0.776	0.180	0.719
Subsidized Credit	0.536	0.197	0.176	0.395	0.193	0.402	0.962	0.461	0.327	0.575	0.566	0.244
Controls												
Interactions	Yes	No	No	Yes	No	No	No	No	No	No	Yes	No
Squares	Yes	No	No	No	No	Yes	Yes	No	No	No	No	No

HARVESTED AREAS	RICE				COFFEE				PLANTAIN			
Matching Procedure	5NNWR	CALIPER			NNWR	CALIPER			NNWR	CALIPER		
		0.2 SD	0.5 SD	1 SD		0.2 SD	0.5 SD	1 SD		0.2 SD	0.5 SD	1 SD
Total Average Effect	0.023	0.021	0.026	0.029	0.029	0.035	0.030	0.030	0.014	0.011	0.010	0.014
P-value	0.029	0.000	0.012	0.014	0.000	0.000	0.000	0.000	0.001	0.000	0.010	0.001
Sample	227	69	169	192	486	208	421	472	574	283	492	551
Treated	113	38	91	103	241	95	197	232	286	138	246	275
Non treated	114	31	78	89	245	113	224	240	288	145	246	276
Balance Test P-Values												
Agric. Potentiality	0.000	0.671	0.486	0.192	0.216	0.845	0.230	0.209	0.120	0.512	0.165	0.862
I&D Potential	0.869	0.478	0.178	0.265	0.884	0.007	0.183	0.209	0.939	0.329	0.222	0.236
Rural Population	0.735	0.932	0.133	0.131	0.967	0.024	0.275	0.497	0.985	0.342	0.477	0.235
Total Population	0.815	0.334	0.122	0.062	0.398	0.007	0.328	0.781	0.661	0.776	0.180	0.719
Subsidized Credit	0.430	0.176	0.395	0.168	0.193	0.118	0.763	0.461	0.327	0.575	0.566	0.244
Controls												
Interactions	No	No	Yes	No	No	No	Yes	No	No	No	Yes	No
Squares	No	No	No	Yes	No	No	No	No	No	No	No	No

Note: SD stands for standard deviation. NNWR stands for nearest neighbor with replacement.

We stress that the parameter size does not vary that much between products (around of 2.3 percentage points for rice and 1.4 for plantain). This because unlike the previous case (the need of irrigation systems), all crops require rural roads to transport inputs and marketing the output, without large distinction between them. Interestingly, the positive and significant impacts of better rural roads coverage on planted areas are more or less of similar sizes to the harvested areas across all crops. Using NNWR models again, the total average effect attain 2.9, 2.5 and 1.4 percentage points for harvested areas of coffee, rice and plantain, respectively (see the lower panel of Table 4).

Finally, results of the role of better access to retail and wholesale markets on planted and harvested areas, are shown in Table 5. Once more we select three crops (coffee, beans and plantain) and the matching procedures that achieved better balance. We note that the better market access facilities are determined by an ample and adequate road network and also by the availability of nearby marketing centers. These two factors, captured through isochrones, could help farmers to be more competitive, since a better infrastructure for marketing could reflect, ultimately, lower transaction costs.

For coffee, the ATT parameter suggests that the planted area in those municipalities with travel times (to the nearest market) below the median is 4 percentage points higher, compared to planted areas in municipalities that are above the median (Caliper 0.2 SD). The sample for this model is 212 municipalities (113 treated and 99 non-treated) and the cross-section estimates are made for 2013. Two ending remarks are taken from Table 5. Firstly, the size of the impact upon planted areas are slightly lower in the other two crops (0.2% and 1.9% for beans and plantain, respectively) and, secondly, there are not substantial differences over the impact size between planted and harvested areas. In fact, the bottom panel of Table 5 shows that the size of the effects are practically equal (especially with caliper 0.2 SD). All these parameters present high statistical significance and the behavior between the sample size and the balance scores tests on controls provide evidence on the robustness of the results.

6. Conclusions

We pursued two objectives in this paper, both of which focused in the role of rural infrastructure on the development of agriculture in Colombia. These are the coverage estimation of both irrigation systems and marketing infrastructure and also the evaluation of the impact of these public goods on the yield of the main crops as well as on their planted and harvested areas. The irrigation and drainage facilities are analyzed because it is expected that permanent water access increases production per hectare and because historically it has been the leading agricultural policy program of land improvement.

Table 5. Impacts of facility to access to markets on Planted and Harvested areas for Some Selected Crops. Estimate by GM

PLANTED AREAS	COFFEE				BEANS				PLANTAIN			
Matching Procedure	NNWR	CALIPER			5NNWR	CALIPER			NNWR	5NNWR	CALIPER	
		0.05 SD	0.1 SD	0.2 SD		0.05 SD	0.1 SD	0.2 SD			0.05 SD	1 SD
Total Average Effect	0.029	0.017	0.026	0.040	0.000	0.007	0.004	0.002	0.020	0.020	0.060	0.019
P-value	0.000	0.000	0.000	0.000	0.999	0.000	0.000	0.012	0.000	0.000	0.000	0.000
Sample	486	21	48	212	543	47	78	205	574	574	22	532
Treated	241	12	26	113	212	22	38	93	206	206	10	193
Non treated	245	9	22	99	331	25	40	112	368	368	12	339
Balance Test P-Values												
Agric. Potentiality	0.919	0.438	0.632	0.643	0.879	0.964	0.376	0.558	0.801	0.845	1.000	0.331
I&D Potential	0.514	0.598	0.005	0.211	0.070	0.219	0.707	0.276	0.005	0.147	1.000	0.410
Rural Population	0.679	0.786	0.337	0.510	0.001	0.545	0.640	0.561	0.478	0.005	0.377	0.107
Total Population	0.445	0.237	0.112	0.001	0.174	0.003	0.002	0.001	0.327	0.143	0.286	0.068
Subsidized Credit	0.400	0.347	0.821	0.451	0.073	0.833	0.964	0.027	0.337	0.102	0.635	0.208
Controls												
Interactions	No	No	No	No	Yes	Yes	No	Yes	No	No	Yes	Yes
Squares	No	No	Yes	No	No	No	Yes	Yes	No	No	Yes	No

HARVESTED AREAS	COFFEE				BEANS				PLANTAIN			
Matching Procedure	NNWR	CALIPER			5NNWR	CALIPER			NNWR	CALIPER		
		0.05 SD	0.1 SD	0.2 SD		0.05 SD	0.1 SD	0.2 SD		0.05 SD	0.1 SD	0.2 SD
Total Average Effect	0.029	0.017	0.024	0.040	0.000	0.010	0.004	0.002	0.020	0.057	0.031	0.022
P-value	0.000	0.000	0.000	0.000	0.844	0.000	0.000	0.012	0.000	0.000	0.000	0.000
Sample	486	21	45	212	543	40	78	205	574	23	97	192
Treated	241	12	24	113	212	19	38	93	206	10	48	92
Non treated	245	9	21	99	331	21	40	112	368	13	49	100
Balance Test P-Values												
Agric. Potentiality	0.919	0.438	0.909	0.643	0.871	0.716	0.376	0.558	0.801	0.318	0.039	0.829
I&D Potential	0.514	0.598	0.001	0.211	0.166	0.430	0.707	0.276	0.005	0.318	0.035	0.965
Rural Population	0.679	0.786	0.154	0.510	0.002	0.895	0.640	0.561	0.478	0.654	0.926	0.273
Total Population	0.445	0.237	0.152	0.001	0.237	0.048	0.002	0.001	0.327	0.243	0.370	0.000
Subsidized Credit	0.400	0.347	0.854	0.451	0.109	0.618	0.964	0.027	0.337	0.569	0.973	0.002
Controls												
Interactions	No	No	Yes	No	No	No	No	Yes	No	No	No	Yes
Squares	No	No	Yes	No	Yes	Yes	Yes	Yes	No	Yes	No	Yes

Note: SD stands for standard deviation. NNWR stands for nearest neighbor with replacement.

We also focused on rural roads, and retail and wholesale centers as the main marketing infrastructure components due to the lower transaction costs that better equipment and maintenance could carry, thus encouraging farmers to intensify their production processes via expanding the agricultural frontier. Since these rural infrastructure goods are public in the sense of non-rivalry and non-exclusion, it was examined the government budget distribution to finance diverse agricultural programs.

The impact of these public goods on agriculture are estimated using genetic matching models for four of the leading crops that are harvested in most of the municipalities: coffee, rice, beans and plantains. The empirical strategy entailed controlling for the most relevant factors, as the degree of agricultural potential (feature that captures the effect of biophysical characteristics on soil fertility); the land potentially apt for irrigation; total and rural population and subsidized credit granted through the public agency Finagro. Additionally, we choose a set of suitable matching algorithms to test the robustness of our findings.

The results confirmed the expected effects. The impact of irrigation and drainage systems on crop yields is positive and statistically significant across all models and crops. The findings support the premise that improving productivity in rural activities requires the provision of public goods like these. The larger impact is on the rice yield (oscillates between 1.7 and 3.5 tons per hectare and year) could be associated with the higher humidity requirements for its production cycle which might be provided by irrigation systems. Our results are coherent with international evidence that has concluded that the yield per hectare of crops benefiting from irrigation systems far exceeds that of those who do not have such schemes, which is reflected finally in large differences in both value of production and total factor productivity.

The impact of infrastructure for marketing upon the planted and harvested areas are no less important. For instance, the planted area in those municipalities with lower travel times to the nearest market (below the median) is 4 percentage points higher, compared to planted areas in municipalities that are further away (above the median). The lower travel times depend mainly of an adequate road network as well as the availability of nearby marketing centers. The planted and harvested areas are also benefited directly from a better network of rural roads.

To the best of our knowledge, in addition to the first attempt made by Lozano and Restrepo (2016), the effects of rural infrastructure over agricultural indicators had not been quantified before this work. We consider that it is the main contribution of this paper. The results are crucial in the practical implementation of public policies. Finally, because the provision of agricultural infrastructure is associated with the size and management of public budget, from the findings of this paper we recommend to reorient agricultural policy towards the supply of public goods that pushes up productivity.

References

Calderón, C., & Servén, L. (2010). Infrastructure in Latin America. *World Bank Policy Research Working Paper Series, Vol*.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, *22*(1), 31-72.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, *84*(1), 151-161.

Diamond, A., & Sekhon, J. S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Review of Economics and Statistics*, *95*(3), 932-945.

Dregne, H. E., & Chou, N. T. (1992). Global desertification dimensions and costs. *Degradation and restoration of arid lands*, 73-92.

Evenson, R. E., Pray, C., & Rosegrant, M. W. (1998). *Agricultural research and productivity growth in India* (Vol. 109). Intl Food Policy Res Inst.

Echavarría, J.J (2014), "El mercado del café en Colombia", in J. J. Echavarría, P. Esguerra, D. McAllister, C. F. Robayo, Misión para la competitividad del café en Colombia.

Fan, S., Zhang, X., & Rao, N. (2004). *Public expenditure, growth, and poverty reduction in rural Uganda* (No. 4). International Food Policy Research Institute (IFPRI).

FAO (1996). Agriculture and Food Security. World Food Summit. Food and agricultural Organization of the United Nations, Rome.

Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, *65*(2), 261-294.

INCODER (2012). La Adecuación de Tierras en Colombia: Antecedentes, Estado Actual y Prospectiva. Mimeo

LaLonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *The American economic review*, 604-620.

Lozano, I., & Restrepo, J. C. (forthcoming). *El Papel de la Infraestructura Rural en el Desarrollo Agrícola en Colombia*. El desarrollo equitativo, competitivo y sostenible del sector agropecuario en Colombia. BANCO DE LA REPÚBLICA.

Mendes, S. M., Teixeira, E. C., & Salvato, M. A. (2008). *Effect of infrastructure investments on total factor productivity (TFP) in Brazilian agriculture* (No. 53441). Universidade Federal de Vicosa, Departamento de Economia Rural.

OECD (2015). COLOMBIA. Policy Priorities for Inclusive Development. Better Policies Series, Enero

Peikes, D. N., Moreno, L., & Orzol, S. M. (2008). Propensity score matching. *The American Statistician*, *62*(3).

Pinstrup-Andersen, P., & Shimokawa, S. (2006). Rural infrastructure and agricultural development. *World Bank*.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Schoengold, K., & Zilberman, D. (2007). The economics of water, irrigation, and development. *Handbook of agricultural economics*, vol *3*, c 29, p 2933-2977.

Shenggen, F. A. N., & Zhang, X. (2004). Infrastructure and regional economic development in rural China. *China economic review*, *15*(2), 203-214.

Smith, J. A., & Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators?. *Journal of econometrics*, *125*(1), 305-353.

Steiner, R., Salazar, N., & Becerra, A. (2015). La política de precios del café en Colombia. Mimeo, Fedesarrollo.

Timmer, C. P. (2002). Agriculture and economic development. *Handbook of agricultural economics*, *2*, 1487-1546.


BANCO DE LA REPÚBLICA
COLOMBIA